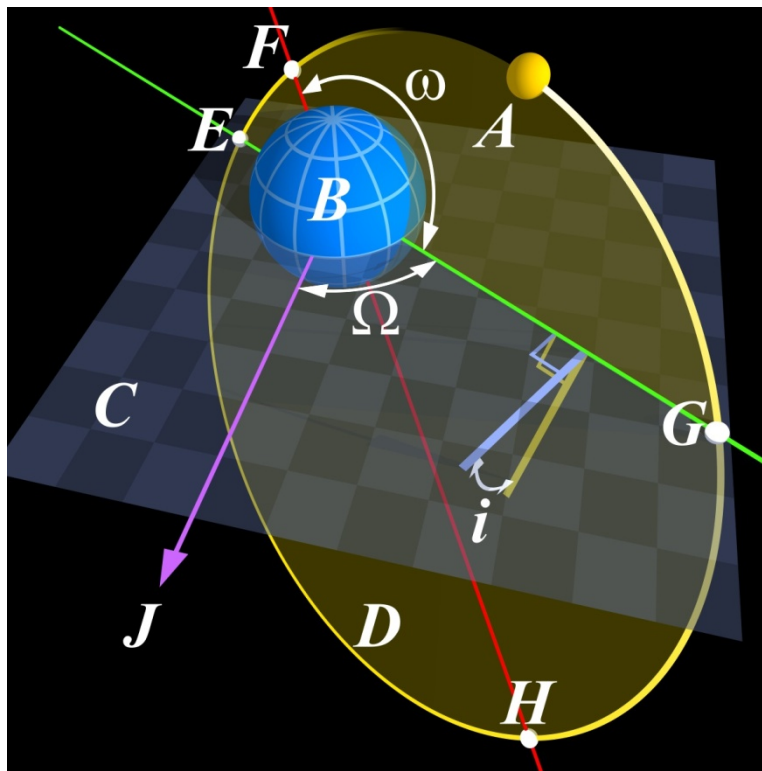


# Celestial Mechanics



Author : J. B. Calvert



---

# Contents

- i. Introduction
- ii. Newton's Laws
- iii. Keplerian Orbits
- iv. Coordinates and Orbits
- v. Cometary Orbits
- vi. Determination of Orbits
- vii. Earth Satellites
- viii. Other Orbit Lore
- ix. References

## Introduction

The grand title for this article conceals the fact that all I want to do here is to show how to use published orbital elements to find the location of a planet in the solar system, and to provide a good explanation for what is involved in the process, including the elements of Newtonian mechanics. I have derived all the results used, with one or two minor exceptions, but it is not necessary to understand the derivations to use the results. However, it certainly helps to know what you are doing.

Johannes Kepler (1571-1630) discovered that the planets moved in elliptical orbits, and his three laws permit the calculation of planetary position once the orbit is known. Although correct, the loss of ideal circular motion as a fundamental (though impossible to apply) concept was a disappointment. This was changed into beautiful triumph in 1686 by Isaac Newton (1642-1727), who showed that the observed

orbital motion was a consequence of fundamental principles of universal application.

Newton's formulation of mechanics, which involved the new concepts of mass and force, was subjected to intense, but sterile, criticism by Ernst Mach (1838-1916) and others, which did not change its application to the slightest degree, and shed no light on its fundamentals. The ideas of Albert Einstein (1879-1955), on the other hand, modified the fundamentals essentially, especially in the theory of gravity and the kinematics of motion, making the theory able to explain the smallest discrepancies with the earlier predictions. This *theory of relativity* has also been abundantly proved by observation. Except for Newton and Einstein, all the vast amount of work in mechanics has been the elaboration and application of the theory, which has proved in every respect correct. Particularly notable is the work of Joseph Louis Lagrange (1736-1813) in [The Beautiful Theory](#), where forces are replaced by work and energy, and the laws of mechanics can be expressed as problems in maxima and minima. This branch of mechanics is usually called *analytical mechanics*, and mechanics with forces is called

*vector mechanics*, from the convenient use of that formalism.

Newton's theory of motion, in its original form, is still the basis of engineering mechanics, and a fundamental part of engineering training. Indeed, it is still used in celestial mechanics. The only aspect needing relativistic corrections so far has been the time effects in the Global Positioning System. In the solar system, velocities are very comfortably less than the speed of light,  $3 \times 10^8$  m/s, although distances are great. Light requires about 8 minutes to pass from the sun to the earth, and gravity travels at the same speed. However, the sun does not care a great deal about the force exerted on it by the earth, while the earth moves in a static gravitational field, so the effects of the finite speed of gravity are not evident. These effects have, however, recently been measured and verified, so it is known that gravity is not instantaneous. We will not need to consider any relativistic effects in the sequel, leaving these tiny effects to another article.

## Newton's Laws

Newton's laws of motion are: (1) a body under no forces moves uniformly in a straight line; (2) the acceleration of a body is equal to the force acting on it divided by its mass; and (3) the mutual forces acting between two bodies are equal and opposite. Newton stated the laws in this form so that these *axioms* would use only common notions and not involve any results following from them, in the spirit of Euclidean geometry. I have used different words in an attempt at conciseness, but the import is the same.

The first law gives an operational method for determining an *inertial system*, a way of describing the position of a mass point in space as a function of time. In the absence of gravitating masses (such as the earth) it is found that a Euclidean space is such a system, in which a uniform time can be approximated by periodic events (such as the

oscillations in lasers). Strictly (and uselessly) speaking, the definition is circular, but in practice it is clearly possible to establish inertial systems at least approximately, and the results obtained agree perfectly with observation.

The second law defines force as the cause of a time rate of change of velocity, by  $d\mathbf{v}/dt = \mathbf{f}/m$ . Again, the definition is logically faulty, but can be realized with great accuracy in practice. Whatever causes a change in velocity in an inertial system is a force. Masses can be compared dynamically, or more easily by the gravitational forces on them. A consistent array of masses and forces can be assembled, which always gives the observed motion. When the velocity is known, the position can be found from the kinematical relation  $d\mathbf{r}/dt = \mathbf{v}$ . Relations not involving force we call *kinematical*, those involving force, *dynamical*.

The third law guarantees that the total linear momentum and total angular momentum will be conserved; that is, that they will remain constant during the motion. These quantities are defined later, and their conservation is the motive for the third law, which cannot be stated in a general form in other



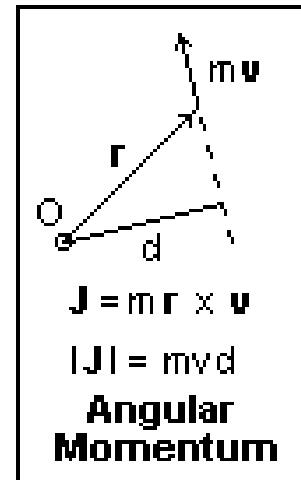
than an extremely tedious and unilluminating way. Of course, not all forces obey it (magnetic forces between currents are an example), but nevertheless momenta are conserved. The conservation of momentum is an important fundamental of mechanics; the third law is a way of introducing it in terms of forces. In celestial mechanics, all the forces are central (act along the line joining two interacting masses) and so the simple statement is sufficient.

The path traced by the tip of the velocity vector  $\mathbf{v}$  as time elapses is called the *hodograph*, and the path traced by the tip of the position vector  $\mathbf{r}$  is called the *trajectory*. An orbit is a special kind of trajectory, a closed (or almost closed) curve. For more information on the hodograph, see [The Hodograph](#). These curves are useful in describing, visualizing and analyzing the motion.

The *momentum* is the product  $\mathbf{p} = m\mathbf{v}$ . The second law can be written  $d\mathbf{p}/dt = \mathbf{f}$ , so that force is the rate of change of momentum. If two particles  $m$  and  $m'$  interact, then  $\mathbf{f}' = -\mathbf{f}$  from the third law, so that  $(d/dt)(\mathbf{p} + \mathbf{p}') = \mathbf{f} + \mathbf{f}' = 0$ , or  $d\mathbf{P} = 0$ , where  $\mathbf{P} = \mathbf{p} + \mathbf{p}'$  is the total momentum. The two particles, taken as a whole, constitute an isolated system, on which no

net force acts, so the rate of change of momentum is zero, and thus the momentum is constant.

The *angular momentum* with respect to a point O is the vector  $\mathbf{J} = \mathbf{r} \times \mathbf{p}$ , where  $\mathbf{r}$  is the radius vector from O to any point on the line of action of  $\mathbf{p} = m\mathbf{v}$ . In the figure,  $\mathbf{J}$  points out of the page. If we use  $\mathbf{r} \times \mathbf{p}$  instead of  $\mathbf{p}$  in the equations of the preceding paragraph, we can prove that the total angular momentum is conserved. The conservation of linear and angular momentum gives us six scalar constants of the motion, useful in analyzing and understanding the motion. The existence of these constants is no trivial matter, but a beautiful and fundamental result of the theory.



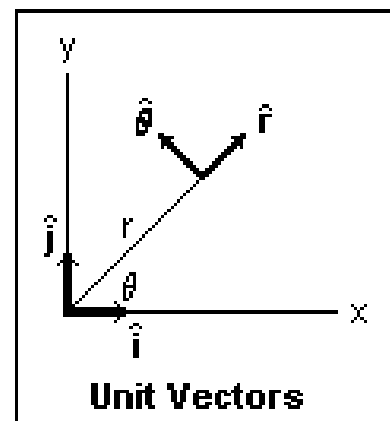
The second law can be formally integrated to give  $\mathbf{p} = \int \mathbf{f} dt = \mathbf{I}$ , where  $\mathbf{I}$  is a vector called the *impulse*, the integral of the force with respect to time. By taking moments about point O, we obtain a similar expression for the angular impulse. Of course, the limits on the integrals must be handled properly.

If the second law is scalar-multiplied by  $\mathbf{v}$ , we find  $m\mathbf{v} \cdot (d\mathbf{v}/dt) = \mathbf{f} \cdot \mathbf{v}$ , or  $(d/dt)(mv^2/2) = \mathbf{f} \cdot \mathbf{v} = -dW/dt$ . This introduces two new scalar quantities, the *kinetic energy*  $T = mv^2/2$ , and the *work*  $-W = \int \mathbf{f} \cdot d\mathbf{r}$ . I am sorry for the minus sign, but it simply means that work is being done *on* the particle, while  $W$  is conventionally the work done *by* the particle. If  $W$  is a function of position, then in vector notation  $\mathbf{f} = -\text{grad } W$ . In one dimension,  $f = -dW/dx$ , which may be more familiar. What we have shown is that  $(d/dt)(T + W) = 0$ , or  $T + W = \text{constant}$ , a new conservation law. If  $W$  is a *conservative* quantity (that is, its value is independent of path and depends only on position), it is called the *potential energy*  $V$ . Then, the *total energy*  $E = T + V$  is conserved in the motion. Energy is a widely used and misused quantity that could be discussed in great detail, but only its bare definition is required here, and the reader is assumed to be familiar with it.

The introduction of energy permits the use of *generalized coordinates*, and the derivation of equations of motion by the Lagrange procedure, which uses the Lagrangian function  $L = T - V$ . This facilitates the solutions of very many problems, since we are liberated from carrying around a basket

of vector components. For more information, follow the link in the Introduction to the article on Lagrangian mechanics.

For vector mechanics, it is convenient to have expressions for the velocity and acceleration components in polar coordinates. These are derived in the References, but will be summarized here.  $\mathbf{v} = (dr/dt)\mathbf{r}' + (r d\theta/dt)\boldsymbol{\theta}'$ .  $\mathbf{r}'$  and  $\boldsymbol{\theta}'$  are unit vectors pointing radially outward, and tangentially in the anticlockwise sense, with respect to the position vector  $\mathbf{r}$ , as shown in the diagram. The rectangular unit vectors  $\mathbf{i}$  and  $\mathbf{j}$  are also shown. Note that  $d\mathbf{r}'/d\theta = \boldsymbol{\theta}'$  and  $d\boldsymbol{\theta}'/d\theta = -\mathbf{r}'$ , while the derivatives with respect to  $r$  are zero.  $d\mathbf{v}/dt = [d^2r/dt^2 - r(d\theta/dt)^2]\mathbf{r}' + (1/r)(d/dt)(r^2 d\theta/dt)\boldsymbol{\theta}'$ .



The reader should carefully note the distinction between, for example,  $d^2r/dt^2$  and  $d^2\mathbf{r}/dt^2$ . The first is the second derivative of a scalar quantity, a simple thing. The second is a vector, and since the direction of  $\mathbf{r}$  can change, the complete expression is somewhat elaborate, as we have just seen. It is essential to distinguish carefully between scalars and vectors, which are arrays of scalars.

These relations are very useful. For example, suppose that the force is central, so that the tangential component of the acceleration is zero. This means that  $(d/dt)(r^2 d\theta/dt) = 0$ , or  $r^2 d\theta/dt = \text{constant} = h$ . Now, half of  $r$  times  $r d\theta$  is the area  $dA$  swept out by the radius vector in  $dt$ , so  $h$  is  $2dA/dt$ . Also,  $mr d\theta/dt$  is the component of the momentum perpendicular to the radius vector, so  $mr^2 d\theta/dt = mh$  is the angular momentum. We then have that the angular momentum  $L = mh$  is a constant. Also, this means that the motion takes place in a plane, and  $L$  is perpendicular to this plane. We have found out a lot here with very little effort. The constancy of areal velocity is one of the most useful properties of orbital motion, not just a curiosity. It relates the rate of change of the angle  $\theta$  to the radius  $r$ .

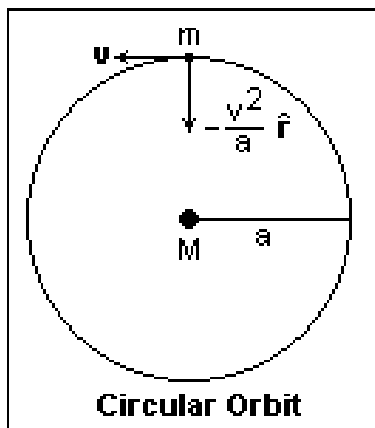
## Keplerian Orbits

Kepler's three laws of planetary motion, deduced by prolonged and tedious consideration of the observed position of Mars, are: (1) the planets move in ellipses with the sun at one focus; (2) the areas swept out by the radius vector in equal time intervals are equal; and (3) The cubes of the mean distances (half the major axis of the orbit) are proportional to the squares of the periodic times. These laws are sufficient to determine the position of a planet at any later time if its position is known at one time, and the dimensions and orientation of the orbit are known.

For a deeper understanding, and the power to attack an arbitrary problem in orbital motion, such as the movement of earth satellites, we should consider the dynamics, on the basis of Newton's theory. The force acting between two spherical, radially symmetric bodies of masses  $M$  and  $m$  a distance  $r$  apart is, by Newtonian gravitation,  $GMm/r^2$  in

magnitude, and is directed along the line joining the centres of the two bodies.  $G$  is the Newtonian gravitational constant,  $6.67259 \times 10^{-11} \text{ m}^3/\text{kg}\cdot\text{s}$ .

Since the force is central, the angular momentum is conserved, and the bodies revolve about one another in a plane with the centre of mass (the barycentre) fixed. The motion can be analyzed as the rotation of a *reduced mass*  $Mm/(M + m)$  about the centre of gravity. The potential energy  $V$  is  $-GMm/r$ , and is negative because the particles attract. For the total energy to be zero,  $T = mv^2/2 = GMm/r$ , or  $v = \sqrt{2GM/r}$ . This is the *escape velocity* from the particle of mass  $M$  at the distance  $r$ . At the surface of the earth, the escape velocity is 11,200 m/s or 25,200 mph. This is also the velocity of a meteor that comes in from infinity. The escape velocity from the *sun* at the distance of the earth is 42,000 m/s or about



95,000 mph. It is not easy to get away from the earth; it is even harder to escape from the sun.

If we consider a circular orbit of radius  $a$ , as shown in the figure at the left, then  $GMm/a^2 = [Mm/(M + m)]a(d\theta/dt)^2$ , since  $v = a(d\theta/dt)$  and the centripetal acceleration is  $v^2/a$ . This is the equation

of motion for the relative motion, with  $r$  the distance between  $M$  and  $m$ . The center of gravity of  $M$  and  $m$  moves with constant velocity, not affected by the relative motion. Thus,  $G(M + m) = a^3(d\theta/dt)^2 = 4\pi^2(a^3/P^2)$ , where  $P$  is the period of the motion. This is a derivation of Kepler's third law for a special case, but it is a general result when  $a$  is the mean distance of any orbit. The quantity  $G(M + m)$  is usually called  $\mu$ , so that  $r^3/P^2 = \mu/4\pi^2$ . If  $r$  is in astronomical units (radius of the earth's orbit, about  $150 \times 10^6$  km), and the period  $P$  is in years, then  $G(M + m) = 4\pi^2 = \mu$ . The quantity  $GM$  is called the Gaussian gravitational constant  $k^2$  if time is in days instead of years, so we get approximately  $k^2 = 4\pi^2/(365.25)^2$ , and  $k = 0.017202$ . A more precise value is 0.01720209895. Untangling  $G$  from  $GM$  means knowing the mass of the sun accurately, which is impossible, so celestial mechanics uses  $k$  instead of  $G$ . The value should be corrected for the finite mass of the earth, but since the ratio of the sun's to the earth's mass is 332,946 the correction is small.

The polar equation of an ellipse, with the origin in a focus, is  $r = p/(1 + e \cos \theta)$ . The *eccentricity* of the ellipse is  $e$ , and  $p$  is a length called the semi-latus-



rectum. In connection with orbits, the angle  $\theta$  is called the *true anomaly*. In an elliptical orbit,  $0 \leq e < 1$ . When  $e = 0$ , we have a circular orbit  $r = p$ . The same formula gives a parabola for  $e = 1$ , and a hyperbola for  $e > 1$ . In an ellipse, the maximum and minimum radii are  $p/(1 - e)$  and  $p/(1 + e)$ , respectively, and the corresponding points on the orbit are called *aphelion* and *perihelion*, respectively. The distance from perihelion to aphelion is the major axis of the ellipse, twice the mean distance  $a$ . Therefore,  $2a = p/(1 - e) + p/(1 + e) = 2p/(1 - e^2)$ , or  $p = a(1 - e^2)$ . The perihelion distance is  $a(1 - e)$ , the aphelion distance  $a(1 + e)$ . The distance from the center to a focus is  $a - a(1 - e) = ea = c$ . This defines the eccentricity as  $e = c/a$ . Since the ellipse is the locus of points the sum of whose distances to the two foci is a constant,  $2a$ , the length of the radius from a focus to the end of the minor axis is  $b$ . Since this is the hypotenuse of a right triangle,  $a^2 = c^2 + b^2$ , where  $b$  is the semiminor axis. Hence  $b = a \sqrt{1 - e^2}$ . The reader can show that the average value of  $r$  over one revolution is  $a$ . This is about all we need to know about the geometry of the ellipse for present purposes. More information can be found in [The Ellipse](#).

If we know the orbit is an ellipse, and the areal velocity is constant, we should be able to prove that the force is central and varies as the inverse square of the distance. On the other hand, if we know that the force is central and inverse-square, then we should be able to prove that the orbit is an ellipse and the areal velocity is constant. Either of these things can be done fairly easily. In the first case, differentiate the expression for  $r$  with respect to time, obtaining  $dr/dt = (e/p) \sin \theta \cdot 2A$ , where  $A$  is the areal velocity (this was called  $dA/dt$  above; now  $A$  represents the derivative). Then differentiate again to find  $d^2r/dt^2 = (4A^2 e/pr^2) \cos \theta$ . The tangential acceleration is zero from the constancy of the areal velocity. The radial acceleration is  $a_r = (4A^2 e/pr^2) \cos \theta - 4A^2/r^3 = -4A^2/pr^2$ , which shows that the force is inverse square, and we have already shown that it is central. In fact,  $GM = 4A^2/p$ .

In the second case, the equations of motion in the plane of the orbit are  $d^2r/dt^2 - r(d\theta/dt)^2 = -GM/r^2$  and  $r d^2\theta/dt^2 + 2(dr/dt)(d\theta/dt) = 0$ . Now,  $(d/dt)(r^2 d\theta/dt) = 2r(dr/dt)(d\theta/dt) + r^2 d^2\theta/dt^2 = 0$ , so the areal velocity  $A = (1/2)r^2(d\theta/dt) = \text{constant}$ . Using this result, the radial equation becomes  $d^2r/d\theta^2 - (2/r)(dr/d\theta)^2 - r = -r^2 GM/4A^2$ . Making the substitution  $u = 1/r$ , this

equation can be thrown into the form  $d^2u/d\theta^2 + u = GM/4A^2$ . This second-order linear equation with constant coefficients is easy to solve. We find  $u = 1/r = GM/4A^2 + C \cos \theta$ . This is the equation of a conic section, as we have seen above. If  $p = 4A^2/GM$  and  $e = pC$ , we have  $r = p/(1 + e \cos \theta)$ , which is what we wanted to prove.

Now we can find the components of the velocity. The transverse component is  $rd\theta/dt = 2A/r = 2A(1 + e \cos \theta)/p$ , where  $A$  is the areal velocity. The radial component is  $dr/dt = 2A(e/p)\sin \theta$ . The square of the velocity is the sum of the squares, or  $v^2 = 4A^2[(1 + e^2 + 2e \cos \theta)/p^2] = 4A^2[(1 - e^2)/p^2 + 2/pr] = -4A^2/pa + 4A^2/pr$ , from which the kinetic energy per unit mass (i.e., setting  $m = 1$  for simplicity) is:  $v^2/2 = -GM/2a + GM/r$ , since  $4A^2/p = GM$ . Rearranging,  $v^2/2 - GM/r = -GM/2a$ , which is to say,  $T + V = E$ , with  $T = v^2/2$ ,  $V = -GM/r$  and  $E = -GM/2a$ . The total energy is negative (i.e., the orbiting particle is "trapped") and a function of the mean distance  $a$  only. The eccentricity is determined by the angular momentum  $h$  through  $e^2 = 1 - h^2/aGM$ . When  $h = \sqrt{aGM}$ , the orbit is circular.

The quantity  $a$ , called the "mean distance," is not the average value of the radius vector  $r$ . It is the average of the perihelion and aphelion distances, however.

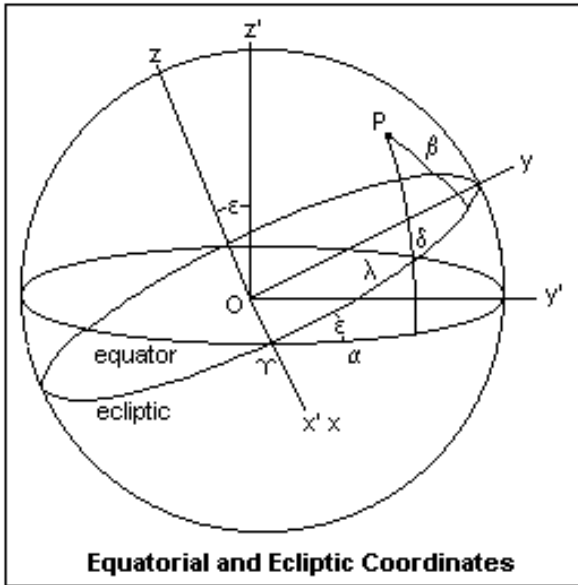
The time average value of  $1/r$  turns out to be  $1/a$ , which is its real definition. A proof will not be supplied here until I find a simple one. This means that the time average potential energy is  $-GM/a$ , and the time average kinetic energy is  $-GM/2a + GM/a = +GM/2a$ . For a circular orbit, the kinetic and potential energies are constant. The time average value of  $r$  is  $a(1 + e^2/2)$ , and the average of  $r$  over the true anomaly is  $b$ , the minor axis of the orbit. The average of  $1/r$  over the true anomaly is  $1/p$ .

We obtained the relation between the mean distance  $a$  and the orbital period  $P$  for the special case of a circular orbit. A more general proof is as follows.

The areal velocity  $dA/dt = h/2 = \text{constant}$ , so integrating from  $t = 0$  to  $t = P$  we find  $A = hP/2 = \pi ab$ . We also know that  $p = a(1 - e^2) = h^2/GM$ .

Therefore,  $P^2 = 4\pi^2 a^2 b^2 / h^2 = [4\pi^2 / GM] a^3$ , since  $b^2/p = a$ . This is the desired general result. If we know the period and mean distance of any orbit, we can calculate the gravitational constant  $GM$ , which in the general case is  $G(M + m)$ .

## Coordinates and Orbits



The angular coordinates that are used to describe directions in space are shown in the figure at the left. The equator and ecliptic planes are shown. The dihedral angle between these two planes is  $\varepsilon = 23^\circ 26' 21''.412$ , or thereabouts. The equatorial coordinates are the

right ascension  $\alpha$  and the declination  $\delta$ , while the analogous ecliptic coordinates are the longitude  $\lambda$  and the latitude  $\beta$ . The relations between the coordinates are best found by considering the rectangular components of the vector  $OP$ , considered as of unit length, and then performing the rotation about the  $x$ -axis that takes one system into the other. We will not require these relations at the present time.

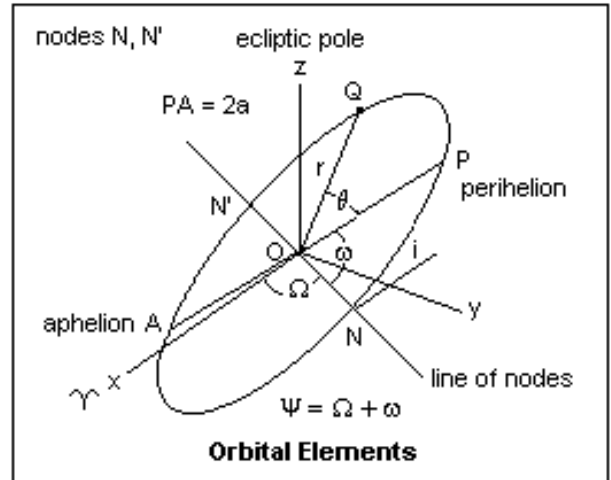
The position of a body in the solar system can be specified in terms of a gigantic rectangular coordinate system centered on the sun. The  $x$ -axis is

always taken as pointing to the vernal equinox, a direction specified by the line of intersection of the orbital plane of the earth (the ecliptic plane) with a plane parallel to the equator of the earth, the equatorial plane. Both contain the centre of the sun. This is made interesting by the precession of the normal to the equatorial plane around the pole of the ecliptic. In the diagram, this is the movement of  $z'$  (the pole of the equator) around  $z$  (the pole of the ecliptic). Fortunately, this is not too rapid, taking about 26,500 years for one revolution. The equatorial plane has to be considered because our earth-based measurements use it as a reference, when directions are given in terms of right ascension and declination. If we know a vector whose components are referred to this plane, we can find the right ascension and declination of its direction. When the earth is located in its orbital plane on the line of intersection of the ecliptic and equatorial planes, the sun is seen in front of the vernal equinox about March 21, climbing from below the equator to above it. At this time, when we look at the sun (careful!) we are looking in the direction of the  $x$ -axis of the heliocentric coordinate system.

The y-axis of the ecliptic coordinates is in the plane of the ecliptic, at right angles to the x-axis, and in the direction in which the planets move, in which a screw would advance in the direction of the north pole of the earth. The ecliptic z-axis is then perpendicular to the x- and y-axes, and makes a right-handed coordinate system with them. The equatorial rectangular coordinates are defined analogously. If we represent these coordinates by primes, x and x' are the same, while x',y' is rotated with respect to x,y by the dihedral angle between the planes. The rotation from the z' axis to the z axis is a right-handed rotation about the positive x,x' axis.

Longitudes are measured clockwise (eastwards) on the ecliptic plane from the vernal equinox, from  $0^\circ$  to  $360^\circ$ . The *longitude of perihelion*,  $\psi$ , is the angle measured clockwise from the vernal equinox to the radius passing through perihelion. (A different symbol, a kind of  $\omega$ , is used conventionally, but it is not available here in HTML.) The longitude of perihelion of the earth's orbit is currently about  $103^\circ$ . This locates the major axis of the orbit and the direction from which the true anomaly is measured.

Other planets move on orbital planes inclined to the ecliptic by the *inclination*,  $i$ . They intersect the ecliptic on a line called the *line of nodes*. The *longitude of the ascending node*,  $\Omega$ , is the angle measured clockwise

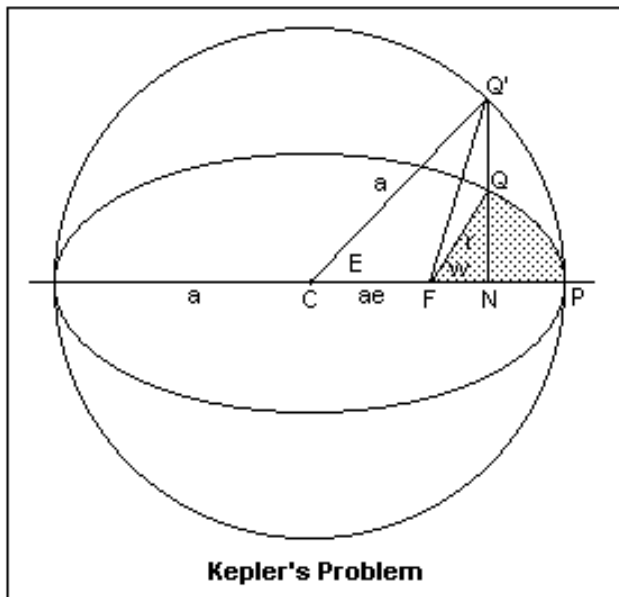


from the vernal equinox to the line of nodes, on the end where the orbit is climbing above the ecliptic plane, called the *ascending node*. The node  $180^\circ$  away is, not surprisingly, the *descending node*. Note in the figure that the traditional symbol for the ascending node is used, which we represent in text by  $\Omega$ , which looks similar. The symbol for the descending node is inverted, with the two little loops on the top. The longitude of perihelion in this case is the longitude of the ascending node, *plus* the angle  $\omega$  in the orbital plane from the node to the perihelion. This may be a little odd, but it does locate the perihelion. That is,  $\psi = \Omega + \omega$ . These quantities are illustrated in the diagram.

Now that the orbital plane and the direction of the perihelion have been specified by the three orbital elements  $i$ ,  $\Omega$  and  $\psi$ , we proceed to specify the shape



of the orbit by its mean distance  $a$  and its eccentricity  $e$ . The position of the body in its orbit at a specified time is given by a sixth element, the mean longitude at the epoch. The "epoch" is just the reference time assumed, usually by its Julian date. The mean distance  $a$  determines the orbital period  $P$  through Kepler's third law. Therefore, only one of  $a$  or  $P$  need be specified. The *mean daily motion*,  $360^\circ$  divided by the orbital period, is often given in place of  $P$ . Orbital elements for planets, asteroids and comets can be found online at the Jet Propulsion Laboratory website whose link is given in the references.



The constancy of the areal velocity is used to locate the body in its orbit at a specified time. This is called Kepler's Problem, which Kepler also solved. The solution is illustrated in the figure. A reference circle of radius  $a$  is drawn around the orbital ellipse. Point  $Q$  is

projected onto the auxiliary circle by a perpendicular to the major axis to point  $Q'$ . The line  $CQ'$  then

defines the angle  $E$  at the centre of the circle, called the *eccentric anomaly*. We can show that it is related to the radius vector  $r$  and the true anomaly  $\theta$  by  $r = a(1 - e \cos E)$  and  $\tan(\theta/2) = \sqrt{[(1 + e)/(1 - e)]} \tan E/2$ . The shaded sector  $PFQ$  is the area swept out by the radius vector, and the area is proportional to the time. The area  $FPQ'$  is proportional to the area  $FPQ$ , since it is merely magnified vertically in the ratio  $a/b = 1/\sqrt{1 - e^2}$ . Indeed, the total area of the orbit is  $\pi ab$ , while the total area of the circle is  $\pi a^2$ .

Therefore, this larger area  $FPQ'$  increases uniformly with time as well. Since the total area is  $\pi a^2$ , the area swept out is given by  $(\pi a^2/P)t$ , where  $t$  is the time since perihelion passage. This area can also be found in terms of  $E$ , since it is the area of the whole sector  $PCQ'$  less the area of the triangle  $CQ'F$ . The area of the sector is  $\pi a^2(E/2\pi) = (a^2/2)E$ . The area of the triangle is  $(1/2)(a \sin E)(ae) = (a^2/2)e \sin E$ .

Therefore, we have  $E - e \sin E = 2\pi t/P = M$ , called Kepler's Equation.  $M$  is called the *mean anomaly* and increases proportionally to time after perihelion passage. This is really a beautiful result, allowing us to find the position in orbit in terms of a uniformly increasing quantity.

Kepler's Equation can be easily solved by iteration. If it is written  $E = M + e \sin E$ , we find the iteration equation  $E_{n+1} = M + e \sin E_n$  ( $e$  is the eccentricity here, not the base of natural logarithms). Starting with  $E_0 = M$ , a few iterations are enough to get good precision, especially when  $e$  is small. The iterations are very easy, and can be performed on a pocket calculator. For  $e = 0.1$ , six iterations give 8 digits. The Newton-Raphson method can also be used. This method finds the roots of the equation  $f(x) = 0$  by improving an initial guess  $x_0$  by the formula  $x_{n+1} = x_0 - f(x_n)/f'(x_n)$ . A sketch will reveal the reasoning behind this formula. In this case,  $f(E) = M - E + e \sin E = 0$  is the function, and  $f'(E) = e \cos E - 1$ . Fewer iterations are required, but each requires more work, than in the case of simple iteration. In practice, an expansion in powers of  $e$  is often used for small  $e$ , but the iteration methods are no more trouble, especially for a computer.

Although the determination of position in orbit may sound complicated, it is really quite straightforward, and after you have done it a few times, it will be easy. Let's find the position of the earth on 2003 April 2. The figures come from the *Astronomical Almanac for 2003*, p. E4. On Jan 1, JD 245 2640.5,

the mean longitude of the earth was  $100.2440^\circ$ , and the longitude of perihelion was  $103.0^\circ$ . April 2 is 91 days later, so  $M = 100.2440 - 103.0 + 91 \times 0.9856 = 86.93^\circ = 1.5172$  radians. A first approximation to  $E$  is  $M + e \sin M$ , and  $e = 0.0167$ , so the correction is  $0.01668$  radians, and  $E = 1.5339$  radians or  $87.89^\circ$ . There is no need to iterate again. The mean distance can be taken as  $1.00000$ , so  $r = 0.9994$  and  $\theta = 88.84^\circ$ . The earth's longitude is  $103.0 + 88.84 = 191.84^\circ$ . On p. C8, the solar ephemeris gives longitude  $11.8506^\circ$  and  $r = 0.9994$ . The solar longitude is  $180^\circ$  less than the earth's, or  $11.84^\circ$  according to our calculation. In both cases, we are using the mean ecliptic of date, and came quite close using Keplerian orbital elements instead of the more elaborate calculations used to calculate the solar ephemeris. Precise orbits require the consideration of planetary perturbations, and this is a very difficult subject indeed.

The orbit of the moon is an example, since it is subject to perturbations by the earth's spheroidal shape and the sun that make its orbit vary in a very complicated way. In the case of the earth's orbit, we really calculate the position of the earth-moon barycentre. This point is  $4671$  km from the centre of

the earth, about two-thirds of the way to the surface. The mean distance of the moon is  $3.844 \times 10^5$  km, and its sidereal period is 27.321661 days, from which  $G(M + m)$  can be calculated and used to relate mean distance and period for other earth satellites after correcting for the mass of the moon. The moon's radius is 1738 km, and its mass  $7.3483 \times 10^{22}$  kg.  $m/M = 0.0123$ , which is not quite negligible. The eccentricity of the moon's orbit is about 0.0549, and its inclination  $5.145396^\circ$ . The plane of the orbit and the position of perigee change rather rapidly. The small eccentricity of the earth's orbit suggests that the earth-moon system did not suffer any external disturbance during its formation, and in particular any collision which created the moon. This was a one-time event, so speculations on what happened can never be proved, since no similar state is known, or is ever likely to become known.

Except for Mercury and Pluto, orbital inclinations are less than  $4^\circ$ , and eccentricities less than 0.0936, the value for Mars. Kepler was fortunate to have picked Mars for study because of its relatively large eccentricity, about twice those of Jupiter and Saturn. Venus's orbit is practically circular, with  $e = 0.0067$ . Mercury not only has a large inclination,  $7.005^\circ$ , but

an eccentric orbit, with  $e = 0.2056$ . Only Pluto has a greater inclination and eccentricity, but it is probably a special case. Observations of Mercury and Pluto were not available to Kepler, who had to work with Tycho's naked-eye observations.

The angle made by  $r$  with the  $x$ -axis in our problem was  $191.84^\circ$ , so the coordinates will be  $x = 0.9994 \cos 191.84^\circ = -0.97814$ , and  $y = 0.9994 \sin 191.84^\circ = -0.20506$ , in astronomical units. In the more general case of a planet in an inclined orbital plane, the radius is first projected on the line of nodes [ $a = r \cos (\omega + \theta)$ ] and on a perpendicular (dip) line [ $b = r \sin (\omega + \theta)$ ]. Then  $z = b \sin i$ ,  $y = b \cos i \cos \Omega + a \sin \Omega$ ,  $x = a \cos \Omega - b \cos i \sin \Omega$ . It is just a matter of projecting  $r$  on the coordinate axes. The necessary formulas are given on page E4 of the almanac. There is an expansion for the true anomaly directly in terms of the mean anomaly, without going through the eccentric anomaly.

To find the equatorial rectangular components, rotate about the  $x$ -axis through an angle of  $i$ . This gives  $y' = y \cos i - z \sin i$  and  $z' = y \sin i + z \cos i$ . If  $x'$ ,  $y'$  and  $z'$  are the components of a vector in equatorial rectangular components, then the right

ascension  $\alpha$  is given by  $\cos \alpha = x'/\sqrt{x'^2 + y'^2}$ , and the declination  $\delta$  by  $\sin \delta = z'/\sqrt{x'^2 + y'^2 + z'^2}$ . The proper quadrants have to be determined. In this way, the direction in which a planet is seen from the earth can be found.

## Cometary Orbits

Comets can be divided into two classes, the *short-period* comets, like Halley's Comet, with a period of 76 years, and *long-period* comets, like Hale-Bopp (1995 O1), which do not return for thousands of years, if at all. A period of 200 years is the conventional dividing-line. Short-period comets have elliptical orbits like the planets, except that the eccentricity is larger, the inclination can take any value, and the comets can move in a *direct* or *retrograde* direction. Retrograde motion is usually expressed by an inclination greater than  $90^\circ$ . Comets were the first real test for Newton's theory, which finally showed that they were normal members of the solar system, not mysterious atmospheric happenings, as had always been supposed. Positions of short-period comets are calculated in the same way as planetary positions, and the orbital elements are presented the same way.



Long-period comets come from the periphery of the solar system, where they wander in the hypothetical Oort Cloud of cometary debris, normally water and carbon dioxide ice and dust. Their total energy is about zero, so the eccentricity of their orbits when they make an excursion towards the sun is about 1. That is, their orbits are parabolic. Short-period comets are those that have suffered an energy-losing collision with a planet (usually Jupiter) and have dropped into an elliptical orbit. There must also have been comets that have gained energy, entered hyperbolic orbits, and were ejected from the solar system. Few, if any, comets have eccentricities significantly greater than 1, which would indicate that they were encountered by the solar system in its path through space, and are not part of the family.

The parabola is a considerably simpler orbit than the ellipse. Its polar formula is  $r = p/(1 + \cos \theta) = (p/2) \sec^2 (\theta/2)$ . The areal velocity  $A = \sqrt{GMp/4}$ . Setting  $x = \tan (\theta/2)$ , direct integration gives Kepler's Equation as  $x + x^3/3 = \sqrt{(GM/2)(2/p)^{3/2}}(t - T)$ . The one real root of this equation gives the position in orbit. There are many ways to solve a cubic equation: refer to mathematics handbooks (see References), algebra texts or mathematics programs.

The time to move from perihelion to the end of the latus rectum is  $t_1 = (1/2)\sqrt{(p^3/GM)}$ .  $2t_1$  is a good measure of the time that the comet will spend near the sun.

The orbital elements include the inclination  $i$ , longitude of the ascending node  $\Omega$ , and argument of perihelion  $\omega$ . The perihelion distance  $p/2$  is given, and the JD of perihelion passage. For Hale-Bopp, the perihelion distance was 0.91399384 and the time of perihelion passage was JD 245 0539.60742. The inclination was  $89.42064850^\circ$ , longitude of ascending node  $282.47215310^\circ$ , and argument of perihelion  $130.59561740$ . From these elements, it was possible to find the position of the comet as seen from the earth. Cometary orbital elements are available on the internet soon after the discovery of a comet. The Jet Propulsion Laboratory supplied the elements for Hale-Bopp shortly after it was discovered.

Halley's comet was the first periodic comet to be recognized. Edmund Halley (1656-1742), friend of Newton's and later Astronomer Royal, suspected that the comets of 1531 and 1607 were the same as the comet of 1682, since they had similar orbits. He

predicted the return of the comet in 1758, which duly occurred. Halley had published Newton's *Principia* at his own expense in 1678, and made an extensive study of comets using the new theory. Halley's comet is the brightest of the short-period comets; most are quite dim, and can even have orbits of small eccentricity like asteroids.

To show how to use the orbital elements given by JPL, let's look at them for Halley's comet. The time of perihelion passage,  $T_p$  is shown as 19860209.45895, which I interpret to mean 1986 February 9, at 0.45895 part of the day, or 11h 0m 53s UT. I should have preferred the Julian Day, which would be unambiguous. The "epoch" is the date to which the elements apply, which is given as 46480. This is a modified JD. The JD can be obtained by adding 24400000.5, or JD 2446480.5, which is 1986 February 18. The size and shape of the orbit is specified by the perihelion distance  $q = 0.58710374$ , in AU, and  $e = 0.96727724$ . The eccentricity is nearly, but not quite, unity. Near perihelion, the orbit will be indistinguishable from a parabola, so the parabolic formulas can be used for predicting its position while it is near the sun and visible. The orientation of the orbital plane is

specified by the longitude of the ascending node,  $58.86004^\circ$ , and the inclination  $i = 162.24220^\circ$ . When the numbers are located, look at the symbols used in the website, which may not be the usual ones because of font limitations. Since  $i > 90^\circ$ , the motion of the comet in its orbit appears to be retrograde, opposite to the direction of motion of the planets, if the inclination is taken to be  $180^\circ - 162.24220^\circ = 17.75780^\circ$ . Finally, the orientation of the orbit in its plane is given by the argument of perihelion,  $111.8656^\circ$ , measured from the ascending node in the direction of movement. It is not easy to comprehend the orbital position from the bare numbers; a drawing, made with some care, will show exactly what is going on.

As the perihelion distance is  $q = a(1 - e)$ , it is a simple matter to determine  $a$ . In fact,  $a = q/(1 - e) = 0.58710374/0.03272276 = 17.941755$  AU. From this, the orbital period  $P$  can be found. Since  $P^2/a^3 = 1$  if  $P$  is in years (earth = 1) and  $a$  is in AU (earth = 1),  $P = 75.99716$  years, very close to 76.0 years. The aphelion distance is  $2(17.941755) - 0.58710374 = 35.296$  AU. This is outside the orbit of Neptune ( $a = 30$ ) but not as far as Pluto ( $a = 39.5$ ). In accordance with the law of areas, Halley's comet spends most of

its time drifting in this dark region, periodically darting in to the sun to have a little more of itself boiled off into space.

Although orbital elements are given to high precision, the accuracy is not necessarily as high. Especially for comets, they can be changed by planetary perturbations, sometimes by large amounts. The ejection of gases from comets also leads to reaction forces that can affect the orbit. Also, the reference coordinates may change with time, because of precession of the equinoxes and other effects, and this must be taken into account in accurate calculations. These changes are not changes in the orbit, of course.

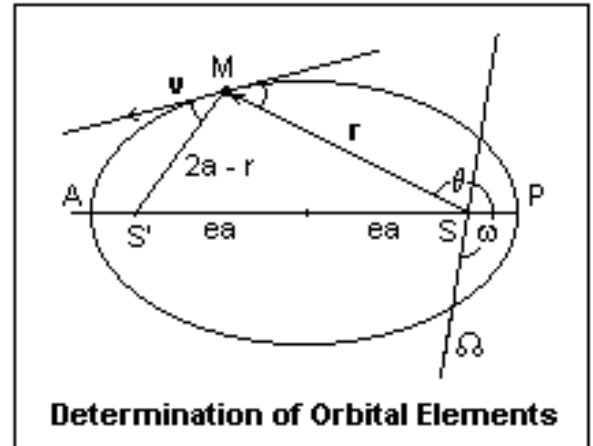
## Determination of Orbits

Finding the orbital elements from observations, and predicting the changes in orbital elements due to perturbations, are two of the most important problems in celestial mechanics, and have received close attention from Newton's time onwards. We cannot give any reasonable account of this work here, but we can show how orbital elements come from observed motions by a graphical analysis that is very instructive, though of little practical use where high precision is required. The reader with drawing supplies is encouraged to follow along.

We suppose known the position  $\mathbf{r}$  of a body  $M$  from the sun and its velocity  $\mathbf{v}$  relative to the sun. These are six parameters that will serve to determine the six orbital elements. The orbital plane is defined by the plane of  $\mathbf{r}$  and  $\mathbf{v}$ , and can be represented in two views by means of orthographic projection (descriptive geometry; see [Monge's Procedure](#)). The line of intersection of the orbital plane with the

ecliptic plane can then be found, and the dihedral angle between the two planes. Since the line of intersection will be the line of nodes, we have now found two of the orbital elements,  $\Omega$  and  $i$ , which determine the plane of the orbit.

The next step is to calculate the mean distance  $a$ , using the energy relation  $v^2 = GM(1/r - 1/2a)$ . From  $a$  we can find the orbital period by  $P^2/a^3 = 4\pi^2/GM$ . The radius vector  $SM$  and the line from  $M$  to the other focus of the ellipse make equal angles with the tangent to the ellipse at  $M$ , which is in the direction of  $\mathbf{v}$ . In a view showing the orbital plane in true size, Lay off  $SM$  and  $MQ$ , where  $MQ$  is an arbitrary line in the direction of the empty focus. The empty focus  $S'$  is located by laying off a distance  $2a - r$  along  $MQ'$ , since the sum of the distances from the foci to  $M$  is  $2a$ . Now we can draw the line  $SS'$ , find its center at  $O$ , and then the locations of perihelion and aphelion. The eccentricity of the orbit is  $e = OS/a = OS'/a$ . The orbit can now be drawn, and the angle from the line of nodes to the perihelion, the argument of perihelion,  $\omega$ , can be measured, as well



as the true anomaly  $\theta$ , which is the angle PSM. We now have the five elements  $\Omega$ ,  $i$ ,  $\omega$ ,  $e$ , and  $a$ .

The remaining element is the time of perihelion passage, which can be found from the true anomaly. First, we find the eccentric anomaly by  $\tan(E/2) = [(1-e)/(1+e)]^{1/2} \tan(\theta/2)$ , and from it the mean anomaly  $M = E - e \sin E$ . The mean anomaly is  $M = 2\pi(t - T)/P$ , where  $T$  is the time of perihelion passage, and  $t$  is the time of observation. That is, the body passed through perihelion at a time  $t - T = MP/2\pi$  earlier. We have now determined all six orbital elements from the six components of the initial position and velocity. This was easy because the distance  $\mathbf{r}$  was one of the given quantities. In general, orbits must be determined by observations of the angular position from earth, not directly in terms of distance and velocity. The position at three different times is sufficient to determine the orbit in this case, but the analysis is more difficult than what we have done.

Orbital elements vary only slowly in the solar system, except when there are near encounters, usually of light bodies with massive planets or satellites, that affect the orbits of the light bodies,



but not of the heavier ones. If we know the orbit at one instant, then we can predict the velocity and position at a later time from the orbit. We can also integrate the actual changes in position and velocity at the later time, taking into account forces exerted by bodies other than the sun. The difference will be reflected in the orbital elements, which will slowly change. This is a useful way to take perturbations into account, and is widely used. The maximum errors in using the Keplerian orbits are on the order of 5" to 30" for the inner planets, somewhat more for Jupiter and Saturn. Neptune was discovered in 1846 as a result of its perturbations of the motion of Uranus. This was a remarkable demonstration of the accuracy of Newtonian mechanics.

The earth's mean distance is decreasing by  $5 \times 10^{-8}$  AU per century, about 75 m per year. The eccentricity is decreasing by  $3.804 \times 10^{-5}$  per century. The inclination is decreasing by 46.94" per century. These are average rates, and are simply the current values of changes that may be periodic. Nevertheless, they show how gradual any changes are, and allow the calculation of the earth's position with reasonable accuracy for thousands of years either side of the present. The longitude of the node,

however, moves at  $-3.038''$  per *year* because of the precession of the equinoxes, which moves the reference point. The argument of perihelion is increasing at  $11.9828''$  per year. It is typical for an orbit to show a relatively rapid change in longitude of the node and motion of the perihelion; the moon is an excellent example, and also earth satellites, which are perturbed by the equatorial bulge of the earth and tidal forces exerted by the sun and moon, as well as by atmospheric drag if they are in low orbits.

## Earth Satellites

The mass of an earth satellite is infinitesimally small compared to the mass of the earth, so the centre of the earth may be considered as a fixed point. A reference frame fixed to the earth's center is not an inertial frame, due to revolution about the sun and the moon's motion. The earth-moon barycentre is a "more inertial" reference point, but in any case the difference from an inertial frame is negligible. There is a change in terminology, as perihelion becomes perigee and aphelion becomes apogee. This useless distinction is regrettable, and can be carried to excess (perijove, etc.). It would be good if there were terms usable in any orbit.

GM for the earth is  $3.98600440 \times 10^{14} \text{ m}^3/\text{s}^2$ .

Therefore,  $a^3/P^2 = 10097$  if  $a$  is in km and  $P$  is in seconds. For  $P = 86164.0989 \text{ s}$  (sidereal rotation period of the earth),  $a = 42,164,172 \text{ m}$ , or an altitude above the surface of 22,232 miles. A satellite with this mean distance will return to the same point in the sky each day. If the inclination and the eccentricity are zero, the satellite will appear to hang

motionless in the sky. Such an orbit is called *geostationary*, obviously of great value as a communications relay location. An area of very choice space real estate is thereby created, for which there is considerable demand. Satellites occupy fixed stations around the equator. They remain in a constant direction from an observing point on the earth's surface, so an antenna can be permanently pointed at them without having to be continually redirected, which would not be easy nor cheap. Their positions may be adjusted slightly to keep the satellites on station by on-board thrusters during their effective life, and the thrusters are used to remove them from orbit when their fuel runs out, since space is at a premium in the geostationary belt.

From a geostationary orbit, the earth subtends an angle of  $17.4^\circ$ , the same as a 12"-diameter globe at a viewing distance of 40", so it is quite reasonable to take photos of the whole visible hemisphere at once. Examples are available at the Dundee University website in the References. The poles are not visible, since the view extends only to latitude  $81.3^\circ$ .

The first artificial earth satellite was Sputnik I, launched on 4 October 1957 and weighing 83.6 kg.

Its orbital period was 96.15 min, so its mean distance was 6953 km. The eccentricity was 0.0517, giving a perigee altitude of 228 km and an apogee altitude of 947 km. The inclination of the orbit was  $64.26^\circ$ , so it could easily be seen in every part of the earth when the sun shone on it. Sputnik remained in orbit until January 1958, making 1350 revolutions. Because of the low perigee, it was considerably affected by atmospheric drag, its final period being near 90 min. The period of a satellite that just grazes the surface of the earth would be 84.48 min, and its velocity 7906 m/s.

The Global Positioning System (GPS) uses satellites that continually radiate very accurate timing information, corrected for relativistic time dilatation due to the satellite's speed, so that their distances from an observer can be found to within a few centimetres from the time differences. The orbits must be known to a similar accuracy, so the effectiveness of this system is evidence of the correctness of the dynamics. Orbit information is broadcast along with timing information. The carrier frequencies are 1.57542 and 1.22760 GHz. The complete GPS system was deployed in 1993, consisting of 21 operational satellites and three

spares, on circular orbits inclined at  $55^\circ$  and with a period of one-half a sidereal day (a sidereal day is about 23h 56m), so the satellites are seen to rise and set about 4 minutes earlier each day, and appear twice a day. Four satellites in good positions are intended to be visible from any point at any time, giving a good intersection, and eliminating the necessity for accurate calibration of the receiver clock (the offset of the receiver clock is one of the four unknowns that can be determined). The altitudes, about 20,200 km ( $a = 26,560$  km), are high enough to make atmospheric drag negligible. There are six orbital planes, with four satellites spaced equally in each. The GPS is by no means the only satellite navigation system that has been developed, but it has become the most used, and is replacing the others.

Earth satellites are affected by many small forces in addition to the main inverse-square force directed toward the centre of the earth, and these *perturbations* cause the orbital elements to change slowly with time. In the case of earth satellites, the mean distance  $a$ , the eccentricity  $e$  and the inclination  $i$  do not change on the average (they may fluctuate slightly over the short term). The longitude

of the line of nodes,  $\Omega$ , the argument of perigee,  $\omega$ , and the rate of change of mean anomaly change steadily with time. The drag of the atmosphere is negligible for satellites that do not come lower than about 1000 km at perigee, which includes most practical satellites. For lower satellites, atmospheric drag causes a loss of energy at perigee that pulls in the apogee position, making the orbit less eccentric and decreasing the orbital period. Paradoxically, the drag speeds up the satellite! The pulls of the moon and sun, *tidal effects* cause orbital changes. The direct tidal effect is due to the force exerted directly on the satellite, while the indirect tidal effect is due to the changes in mass distribution of the earth caused by tidal motions. Pressure of solar radiation and solar wind is another disturbing effect. The most important perturbing force, however, is that exerted by the equatorial bulge of the earth.

The gravitational potential of the earth can be expressed approximately as  $V = GM/r [1 - (a_E/r)^2 J_2 P_2(\sin \varphi)]$ , where  $a_E$  is the equatorial radius of the earth, 6 378 137 m,  $J_2 = 0.001082630$ , and  $P_2(x)$  is the Legendre polynomial of second degree,  $(1/2)(3x^2 - 1)$ . The argument of this polynomial is usually seen as  $\cos \theta$ , but here we use the latitude  $\varphi$

$= 90^\circ - \theta$ . This is the first part of an expansion in spherical harmonics, which can even allow for a lumpy earth that is not symmetric about the polar axis. The correction to the  $1/r$  field shown here is a zonal harmonic of order 2, related to the flattening  $f = a/(a - b) = 1/298$  of the earth. However, the gravitational potential depends on the distribution of mass in the earth, so it cannot be expressed simply in terms of  $f$ . The radial component of the force is  $-\partial V/\partial r$ , and the tangential part is  $(1/r)\partial V/\partial \theta$ . The motion of earth satellites has led to a much better knowledge of the earth's gravitational potential.

If a satellite has a mean motion  $n = 2\pi/P$ , a mean distance  $a$  and an inclination  $i$ , the movement of the line of nodes is  $d\Omega/dt = -(3n/2)(a_E/a)^2[\cos i/(1 - e^2)^2]J_2$ , and the change of the argument of perigee is  $d\omega/dt = +(3n/4)(a_E/a)^2[(5\cos^2 i - 1)/(1 - e^2)^2]J_2$ . The quantities are in MKS units, and the results are in radians per second. For the moon, the line of nodes rotates once in 18.6 years. For a typical GPS satellite, the line of nodes moves about  $-0.03^\circ$  per year, the perigee about  $0.01^\circ$  per year. All the other perturbations are much smaller in amount, but must be taken into consideration in accurate work.

Relativistic perturbing accelerations are inversely



proportional to the fourth power of the distance  $r$ . Their magnitudes for GPS satellites are about  $3 \times 10^{-10} \text{ m/s}^2$ , less than other small perturbations, so they have no practical effect, although they have been considered.

The International Space Station (ISS), shown at the right, is a famous earth satellite. The image is from the NASA website. The first two modules, Zarya and Unity, were assembled in orbit in 1998, and now there are 14. It is 73 m across the solar arrays, 44.5 m long and 27.5 m high, with  $425 \text{ m}^3$  of habitable space. Power comes mainly from the  $892 \text{ m}^2$  of solar arrays. There are thrusters to adjust the orbit when necessary. The total mass is about 179 metric tons. Although this is called "space travel" by NASA, it is not even outside the earth's atmosphere!



Approximate elements of the ISS orbit at 13.49Z, 6 April 2003, are  $i = 51.6^\circ$ ,  $\Omega = 19.99^\circ$ ,  $e = 0.00083$ ,  $\omega = 52.91^\circ$ ,  $M = 307.29^\circ$ , and mean motion  $15.59579861 \text{ rev/day}$ , or  $0.0649825^\circ/\text{s}$ . ("Z" is another designation for UT.) Exact orbits, predicted about 10 days in advance, can be found at the NASA link in the References. The period is  $P = 5539.95 \text{ s}$  (1.54 hr),

and the mean distance is 6767009 m. This is a low, almost circular orbit at an altitude of about 400 km. Because the orbit is low, the ISS is seldom illuminated by the sun at night, so it is not frequently seen, though a frequent, large and prominent object.

The longitude (RA) of the ascending node changes rapidly, by about  $-5.00^\circ$  per day, as does the argument of perihelion, by  $-3.64^\circ$  per day.

Perturbation by the oblateness of the earth is the reason for most of this change. The inclination, eccentricity and mean distance do not change rapidly. The orbit is just above the maximum ionization in the F layer of the ionosphere, and so is affected by atmospheric drag. The density of the atmosphere at this height is about  $\rho = 9 \times 10^{-12}$  kg/m<sup>3</sup>. The drag is given by  $F = C\rho V^2 A/2$ , where  $C$  is the drag coefficient,  $V$  the satellite velocity, and  $A$  the effective projected area. NASA gives  $A = 344$  m<sup>2</sup> and  $C = 2.36$ .  $V = 2\pi a/P = 6541$  m/s (assuming a circular orbit), so  $F = 0.16$  N, which will produce an acceleration of  $8.9 \times 10^{-7}$  m/s<sup>2</sup>. This drag causes the mean distance and period to decrease. NASA notes that the "decay" is  $4.11 \times 10^{-4}$  rev/day<sup>2</sup>. In 100 days, the mean motion will increase by 0.0411 rev/day, or  $P = 5527$  s, a decrease of 13 s, or 0.23%. If

uncorrected, the satellite would spiral inward at an increasing rate, eventually burning up catastrophically.

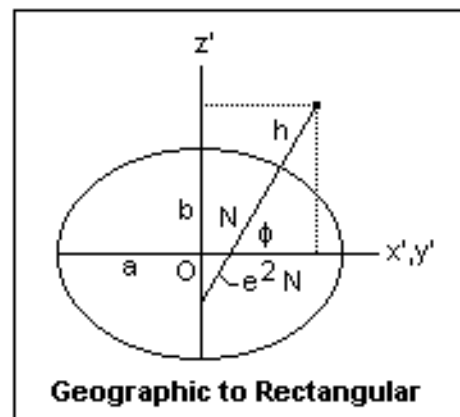
It may be interesting to find out where you should look for a satellite in a known orbit. The procedure will be illustrated here without taking all the refinements into account that are necessary for, say, GPS positioning. The idea is to find the rectangular coordinates of the satellite in an approximately inertial system with its origin at the centre of the earth, and then to find the rectangular coordinates of the point of observation in the same system. The differences of the coordinates then give a vector from the point of observation to the satellite. The motion of a satellite as seen from a fixed location on earth may be very complex, because of the interaction of the two motions involved, the revolution of the satellite and the rotation of the earth. Only when these are approximately equal is the situation more or less simple.

Let's suppose the satellite is in a circular orbit with  $a = 26\,500\,000$  m and  $i = 55^\circ$ , like a GPS satellite. Let the longitude of the ascending node be  $40^\circ$ . Since the orbit is circular, there is no perigee point, so we

measure the true anomaly from the line of nodes in the orbital plane. Let the true anomaly at the time we are considering be  $70^\circ$ . The z-axis is taken along the axis of rotation of the earth, and the x-axis in the direction of the vernal equinox. The y-axis then makes a right-handed coordinate system. This is just like the case of planetary motion, except that the role of the ecliptic plane is played by the equatorial plane. The orientation of these axes remains fixed in space as the earth revolves about the sun.

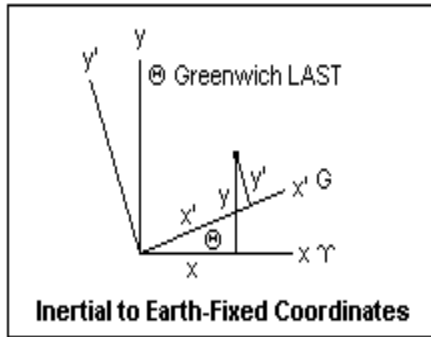
First, resolve the radius vector along and perpendicular to the line of nodes. The components are  $a \cos \theta = 9\,084\,055.0$  m and  $a \sin \theta = 24\,958\,236.0$ . Then,  $z = a \cos \theta \sin i = 20\,444\,590.0$  m,  $x = a \cos \theta \cos \Omega - a \sin \theta \sin \Omega = -9\,084\,055.0$  m, and  $y = a \cos \theta \sin \Omega + a \sin \theta \cos \Omega = 24\,958\,236.0$  m. This is the instantaneous position of the satellite in the inertial system, and it could obviously be found for any time equally easily.

Now for the point of observation. Let's use my own location, which is longitude  $\lambda = -104.92583^\circ$ , latitude  $\phi = 39.72694$  and height above the



ellipsoid (or mean sea level)  $h = 1633.7$  m. Your own coordinates, if you do not know them already, can be found from a USGS topographic map. A meridional section of the earth is shown at the right, where  $\varphi$  is the geographic or spheroidal latitude. The dimensions of the earth are  $a = 6378136$  m and  $b = 6356752$  m. There is a number of reference spheroids, any one of which will work satisfactorily. The eccentricity of the elliptical cross-section of the spheroid can then be found to be  $e = 0.006694167$ . The meridional radius of curvature at latitude  $\varphi$  is  $N = a[\cos^2\varphi + (1 - e^2)\sin^2\varphi]^{-1/2}$ , or  $6\,378\,194.38$  m at my location. The length of  $N$  below the equatorial plane is  $e^2N = 285.8$  m.

We now choose an earth-fixed rectangular coordinate system with the  $z'$ -axis along the rotational axis of the earth, the  $x'$ -axis in the meridian of Greenwich, and the  $y'$ -axis making a right-handed system. The rectangular coordinates of my location are then  $x' = (N + h)\cos\varphi\cos\lambda = -1\,263\,816.21$  m,  $y' = (N + h)\cos\varphi\sin\lambda = -4\,741\,167.76$  m,  $z' = (N + h - e^2N)\sin\varphi = 4\,077\,353.73$  m. These coordinates do not change as the earth rotates. The distance from the centre is  $6\,379\,711.32$  m.



The relation between the inertial and earth-fixed coordinate systems is shown at the left. The earth-fixed system rotates steadily in the positive direction about the common  $z$  and  $z'$  axes, and the angle of rotation  $\Theta$  is the local apparent sidereal time at Greenwich, which is easy to find from the Astronomical Almanac. The rotational period of the earth is 23h 56m 04.09890369732s. When the time is stated to such precision, there are many small corrections to be considered, but we shall ignore them here so the principle is not lost. We shall assume that the inertial and earth-fixed coordinates are related simply by the rotation through an angle  $\Theta$ , not correcting for polar motion, precession of the equinoxes and nutation. These are simply rotation matrices applied in addition to the one for the main rotation. For accurate work, however, they cannot be ignored. Since sidereal time goes from zero to 24 hours in one rotation, we must multiply an ordinary time interval by 24/23.9344719 or 1.00273781195 to find the equivalent sidereal time interval. A sidereal hour is simply equivalent to

15° of angle; it is not an "hour" in the usual sense of time.

The transformation equations are the familiar  $x = x' \cos \Theta - y' \sin \Theta$  and  $y = x' \sin \Theta + y' \cos \Theta$ . All you have to do is check that the signs of the sines are correct. If we put in coordinates in the earth-fixed system  $(x',y',z')$  we will get out coordinates in the inertial system  $(x,y,z)$  which we can compare with the satellite coordinates. Let's pick 10.00 am MST on 2003 January 1. Since my time zone is +7, the UT is 17.00h January 1. From the Astronomical Almanac, page B8, the sidereal time at 0h UT was 6.61651428 hours. The sidereal time elapsed is then  $1.00273781195 \times 17.00 = 17.04654280$  hours. Adding the two, the Greenwich LAST at my 10.00 am will be 23.66305708 hours, or  $354.945856^\circ$ . If it seems easier, this can also be expressed as  $-5.0541438^\circ$ . Now I can find my inertial coordinates at this instant. The results are  $x = -1\,676\,585.46$  m,  $y = -4\,611\,395.05$ ,  $z = 4\,077\,353.73$  m. The square root of the sum of the squares is 6 379 711.32, so the arithmetic checks.

Now we can subtract the coordinates of the observation point from the coordinates of the

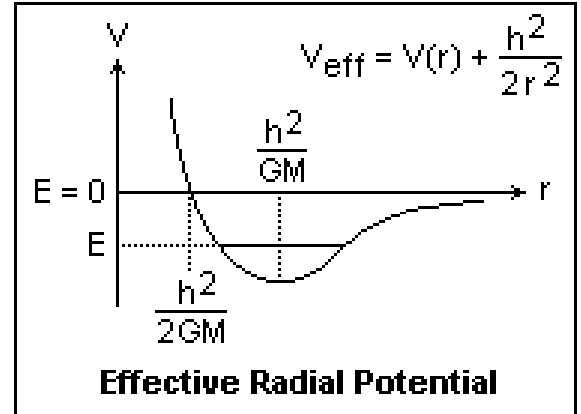
satellite to find the relative vector  $(X,Y,Z)$ . I find  $X = -7\,407\,469.54$  m,  $Y = 29\,569\,631.05$  m,  $Z = 16\,367\,236.27$  m. The direction of this vector can now be expressed in terms of right ascension and declination, since it is a vector in the inertial system referred to the vernal equinox. The distance from the observer to the satellite is  $34\,599\,423.53$  m. The projection on the equatorial plane is  $30\,483\,334.55$  m, so that the declination  $\delta = 28.2324^\circ$  and the right ascension is  $75.9363^\circ$  or  $5\text{h } 3\text{m } 45\text{s}$ . At the date and time specified, this direction is beneath the earth. In fact, the antipodeal point of right ascension  $18\text{h}$  and declination  $-29^\circ$  is in the southern sky at an altitude of about  $20^\circ$  above the horizon.

Although these calculations are tedious and subject to error, even when done with an electronic calculator, a computer program can do them with ease, speed and correctness, and even the more precise calculations are not much bother. It would not be difficult to write a program to determine the visibility of all of the 24 GPS satellites, and the directions in which they are to be seen.



## Other Orbit Lore

The total energy  $E$  of a body of unit mass moving under the influence of a potential energy per unit mass  $V(r)$  is the kinetic energy  $T = v^2/2$  plus  $V(r)$ . In polar coordinates, the



components of velocity are

$dr/dt$  and  $r d\theta/dt$ , so the total energy becomes  $E = (dr/dt)^2/2 + r^2(d\theta/dt)^2/2 + V(r)$ . The constancy of angular momentum for a central force gives  $r(d\theta/dt)^2 = h = \text{constant}$ . Substituting for  $d\theta/dt$ , we find  $E = v^2/2 + h^2/2r^2 + V(r)$ . This is the energy equation for a particle of unit mass moving in one dimension with the *effective* potential  $h^2/2r^2 + V(r)$ . The additional term is called the *centrifugal potential*. In the case of orbital motion under gravity,  $V(r) = -GM/r$ . The form of the resulting potential is shown in the figure. For  $E = 0$ , a particle coming in from infinity accelerates steadily until  $r = h^2/GM$ , then is rapidly decelerated, coming to rest at  $r = h^2/2GM$ , and

afterwards retracing its path in reverse. Of course, the particle swings round the centre of force in this motion. For  $E = -(GM)^2/2h^2$ ,  $r$  has the constant value that gives the minimum of the curve, and the orbit is circular. For intermediate energies,  $r$  oscillates between a minimum and a maximum value, corresponding to perhelion and aphelion.

Consider the vector  $\mathbf{B} = \mathbf{v} \times \mathbf{h} + (GM/r)\mathbf{r}$ . The angular momentum per unit mass,  $h$ , is here taken to be a vector perpendicular to the orbital plane. The cross product then lies in the orbital plane, and so then does  $\mathbf{B}$ . Taking its time derivative, we have  $(d\mathbf{v}/dt) \times \mathbf{h} - (GM/r^3)\mathbf{r}(\mathbf{v} \cdot \mathbf{r}) + (GM/r)\mathbf{v}$ . Now  $d\mathbf{v}/dt = -(GM/r^3)\mathbf{r}$ , and  $\mathbf{h} = \mathbf{r} \times \mathbf{v}$ , so the first term becomes  $-(GM/r^3)[\mathbf{r}(\mathbf{r} \cdot \mathbf{v}) - r^2\mathbf{v}]$ . All the terms cancel, so  $d\mathbf{B}/dt = 0$ , and  $\mathbf{B}$  is a *constant of the motion*. This is extraordinary, since we already have one constant of the motion, the total energy  $E$ , and this is usually the only one that can be found. For an inverse-square force, we have found a second one. It is known that this happens when we have an additional symmetry in the problem--here it happens to be symmetry under rotation in a four-dimensional space, but there is no room to explain this here. We'll just be happy with the result.

Since  $\mathbf{B}$  is a constant, it can be evaluated at any point of the motion, and it is easiest to choose perihelion. At this point,  $\mathbf{v} \times \mathbf{h} = v_p^2 \mathbf{r}_p$ , so  $\mathbf{B}$  is a vector parallel to the line between the focus and perihelion. We will find the same vector if we evaluate  $\mathbf{B}$  at any point of the motion. Its constancy is a good proof of an exact inverse-square force.

We have seen that with an attractive inverse-square force, closed orbits exist for negative energies. For zero or positive energies, the orbits are open. We have considered briefly the zero-energy parabolic orbits that divide the two classes, but not the hyperbolic orbits, since they are not of much use in the solar system. For a repulsive inverse-square force, there are no closed orbits of negative energy at all, only the hyperbolic orbits of positive energy. On these orbits, the body comes in from infinity along a path asymptotic to a line of a certain direction. It then interacts with the other body, in what is called a "collision," though the bodies do not come into contact, and leaves to go to infinity asymptotic to a different direction. The angle between the asymptotes is a function of the *impact parameter*, which is the distance of closest approach if the bodies did not exert forces on each other. This is the

problem of *Coulomb scattering* in atomic and nuclear physics. Rutherford and his students used classical mechanics to analyze the scattering of alpha particles ( $Z = 2$ ,  $A = 4$ ) by atomic nuclei to verify the existence of nuclei. This is an area different enough from celestial mechanics to leave it for another article.

## References

P. van de Kamp, *Elements of Astromechanics* (San Francisco: W. H. Freeman, 1964).

F. R. Moulton, *An Introduction to Celestial Mechanics*, 2nd ed. (New York: MacMillan, 1914).

Y. Ryabov, *An Elementary Survey of Celestial Mechanics* (New York: Dover, 1961).

P. K. Seidelmann, ed., *Explanatory Supplement to the Astronomical Almanac* (Mill Valley, CA: University Science Books, 1992). Chapter 1 is especially recommended.

*The Astronomical Almanac for the Year 2003* (Washington, DC: U.S. Government Printing Office, 2001). pp. B8, C8 and E3-E5. Orbital elements for the minor planets (asteroids) are also included.

Bronshtein and Semendyayev, *A Guide Book to Mathematics* (New York: Springer, 1973). p. 161 (solution of cubic equation).

I. Ridpath, ed., *Norton's 2000.0 Star Atlas and Reference Handbook*, 18th ed. (New York: John Wiley & Sons, 1989). pp. 114-117. There is a table of cometary orbital elements.

L. D. Landau and E. M. Lifshitz, *Mechanics* (Oxford: Pergamon Press, 1960). This is an excellent concise text in classical mechanics, looking towards applications in quantum mechanics and atomic physics.

Orbital elements for planets, asteroids and comets can be found at [JPL](#).

Orbital elements for the ISS can be found at [ISS Orbit](#). The website contains information on NASA missions.

An excellent source of earth satellite orbits is [Dundee University Satellite Station](#). Free registration opens up excellent things, like IR and VIS images from GEOS and other satellites, as well as orbital elements of a number of satellites.

B. Hofmann-Wellenhof, H. Lichtenegger and J. Collins, *GPS Theory and Practice* (Wien: Springer-Verlag, 1993).

